Outlier robust cointegration analysis

Franses, P.H.; Lucas, A.

1997

document version
Early version, also known as pre-print

Link to publication in VU Research Portal

citation for published version (APA)
Franses, P. H., & Lucas, A. (1997). *Outlier robust cointegration analysis*. (Research Memorandum; No. 1997-45). FEWEB.

ulteit der Economische Wetenschappen en Econometrie

erie Research Memoranda

Outlier Robust Cointegration Analysis

Philip Hans Franses
Andre Lucas

Research Memorandum 1997-45

August 1997

vrije Universiteit amsterdam

Outlier Robust Cointegration

Philip Hans Franses
André Lucas

Research Memorandum 1997-45 August 1997

OUTLIER ROBUST COINTEGRATION ANALYSIS

By Philip Hans Franses and André Lucas

Standard unit root tests and cointegration tests are sensitive to atypical events such as outliers and structural breaks. This paper uses outlier robust estimation techniques to reduce the impact of these events on cointegration analysis. As a byproduct of computing the robust estimator, we obtain weights for all observations in the sample. These weights can be used to identify the approximate dates of the atypical events. We evaluate our method using some illustrative simulated data. Furthermore, since our robust approach involves a few additional decisions on the values of key parameters, we investigate the sensitivity of our method through extensive Monte-Carlo simulations. Finally, we present an empirical example based on real-life data to show that OLS-based cointegration tests can spuriously indicate stationarity.

Keywords: Robust estimation, unit roots, cointegration, outliers, structural breaks.

1. INTRODUCTION

Univariate and multivariate unit root tests are sensitive to irregular observations. Perron and Vogelsang (1992), for example, prove that univariate unit root tests are biased towards nonrejection if there is a level shift in the sample. Conversely, unit root tests are biased towards rejection if aberrant observations occur as isolated additive outliers, see Lucas (1995a,b) and Franses and Haldrup (1994) for the univariate and multivariate case, respectively.

Irregular data patterns arise quite naturally in empirical modeling exercises. They are often due to the approximate character of the postulated model. As reality is more complex than the model put forward by the econometrician, it is not surprising that certain observations or periods are not captured adequately by a given model. As such, aberrant observations need not be "bad" observations. They merely reflect the limitations of the model used for describing the data. In particular, aberrant observations can reveal useful information about the working of economic mechanisms and indicate valuable directions for model respecification and/or augmentation. If not accounted for in the appropriate way, however, irregular data patterns may corrupt the results of standard estimation and test procedures completely, see the references mentioned above. Therefore, it seems useful to construct statistical procedures for dealing with aberrant observations in a constructive way. Such procedures must meet two objectives. First, statistical inference on the parameters of interest, in our case the unit root parameters, must not be blurred by a few observations that are not fitted by the model. Second, the statistical procedures must provide a clear signal as to which observations are not captured by the model. These observations can then be subjected to a more thorough analysis in a subsequent step of the modeling process.

One possible solution to the problem of irregular observations is to include zero-one dummy variables in the regression model used to test for unit roots. In a univariate context, for example, Perron (1989) includes of a set of deterministic regressors in order to allow for an alternative hypothesis with a trend break at a known date. The inclusion of such deterministic regressors changes the asymptotic distribution of unit root tests. In a more general setting, the asymptotic distributions of unit root tests change with the presumed location of the irregular observations as well as with the model needed to describe these observations. Additionally, it is very important whether the location of the aberrant observations is known from the outset or not, see, e.g., Christian0 (1992).

In principle, it is possible to extend univariate unit root tests to allow for all possible types of aberrant observations and to develop the relevant asymptotic theory. For practical purposes, however, one then needs considerable skill to entertain all the different models and to evaluate the many different test statistics. Signals as to where the aberrant observations are located can be obtained from outlier detection methods as proposed in, e.g., Chen and Liu (1993) and Tsay (1988). So even in the relatively simple univariate case, it is quite complicated to come up with statistical procedures that statisfy the two objectives formulated earlier.

For multivariate time series, however, the situation is even worse. Extending the approach sketched in the previous paragraph may in that case end in a sheer endless set of decisions to be made by the practitioner. One should decide on the type of models for generating the aberrant observations, and on whether these models should be allowed to differ across equations. Furthermore, one should account for the fact that irregular observations need not occur simultaneously. All these choices can result in a plethora of distinct asymptotic results for a large number of test statistics, thus making the whole approach quite unpractical.

The aim of our paper is to overcome the above problem for the case of a multivariate cointegrated vector autoregressive model. To achieve this, we advocate the use of an estimator that can cope with irregular data. This estimator automatically assigns less weight to aberrant data points. Consequently, the estimator meets the two objectives formulated earlier: aberrant observations have less impact on parameters of interest, and the estimator clearly signals the position of the aberrant observations in the form of small observation weights. These weights may subsequently be used to suggest possible modifications to the model in terms of adding variables and/or allowing for time-varying parameters. An additional advantage of our robust method over the approach based on dummy variables is that our method requires only a single asymptotic theory.

The outline of our paper is as follows. Section 2 discusses the outlier robust cointegration tests and shows how weights can be constructed for the individual observations. Section 3 illustrates the usefulness of our outlier robust method to track specific types of model failure. This is done using some simulated example series. Next, we investigate the sensitivity of our approach to specific values of key parameters that have to be fixed from the outset. Our Monte-Carlo results suggest guidelines for a sensible practical application of our method. Section 4 considers an empirical example concerning the Finland/US real exchange rate. We show that OLS based cointegration tests can be biased towards the alternative hypothesis if additive outliers are present. Section 5 concludes the paper.

2. OUTLIER ROBUST COINTEGRATION TESTS

We begin this section with a discussion of the vector autoregressive (VAR) model and the outlier robust cointegration test in Subsection 2.1. Next, we present the observation weights that follow from robust estimation procedure in Subsection 2.2. These weights can usefully be exploited in a subsequent analysis.

2.1. *The model and test statistic*

We consider the VAR model of order p,

$$\Delta y_t = \alpha\beta' y_{t-1} + \Phi_1 \Delta y_{t-1} + \ldots + \Phi_{p-1}\Delta y_{t-p+1} + \mu + \varepsilon_t, \tag{1}$$

with y_t and ε_t $(k \times 1)$ vector processes, with $\Phi_1, \ldots, \Phi_{p-1}$ $(k \times k)$ parameter matrices, and with α and β $(k \times r)$ parameter matrices of full column rank, where $0 \leq r \leq k$. $\{\varepsilon_t\}$ is assumed to be a white noise process with zero mean and covariance matrix Σ. A denotes the first-difference operator: $\Delta y_t = y_t - y_{t-1}$. Under suitable regularity conditions on the coefficient matrices, one can show that r linear combinations of y_t are stationary, see Johansen (1991). These linear combinations $\beta' y_t$ are called the cointegrating relations, and the columns of β are called the cointegrating vectors. Johansen (1988, 1991) proposes a method to determine the number of cointegrating relations r. Assuming that ε_t in (1) is normally distributed, he derives the likelihood ratio (LR) test statistic for the hypothesis H_r : $\text{rank}(\Pi) \leq r$ versus the alternative H_k : $\text{rank}(\Pi) = k$. The limiting distribution of this test statistic is nonstandard and can be expressed as a functional of Brownian motions.

Franses and Haldrup (1994) show that additive outliers can seriously affect empirical cointegration analysis. In their Table 2 they present the empirical fractiles of the Johansen cointegration test based on many Monte-Carlo replications using time series with various sizes of additive outliers. These fractiles markedly exceed those for the no outlier case. Hence, using the standard critical values in case of outliers leads to spurious cointegration.

In order to reduce the effect of outliers, Lucas (1997) proposes a Johansen-type testing procedure based on non-Gaussian pseudo-likelihoods. The particular implementation of this procedure in the present paper is as follows. We estimate the parameters in (1) using the Student t pseudo-likelihood with ν degrees of freedom:

$$\mathcal{L}(\theta) = \prod_{t=1}^{T} \frac{\Gamma((\nu+k)/2)}{\Gamma(\nu/2)|\pi\nu V|^{1/2}} \left(1 + \frac{\varepsilon_t' V^{-1} \varepsilon_t}{\nu}\right)^{-(\nu+k)/2}, \tag{2}$$

where θ denotes the vector of unknown parameters. Note that we use the likelihood in (2) as a pseudo-likelihood in the sense of Gouriéroux et al. (1984). This means that ε_t need not be Student t distributed itself. In fact, the distribution of ε_t only has to satisfy certain weak conditions, e.g., finite variance, see Lucas (1997). The Student t pseudo-likelihood in (2) is used as a device for mitigating the effect of aberrant data structures on unit root inference. Hoek et al. (1995) show in the univariate case that the Student t pseudo-likelihood can be usefully employed to attain this objective. Lucas (1995b, 1997) furthermore shows that the cointegration test based on (2) has a considerably higher power than the Gaussian-based test if ε_t is leptokurtic. We will demonstrate in Sections 3 and 4 by means of simulations and an empirical example that the result of Hoek et al. (1995) extends to the multivariate case. Moreover, we substantiate the claim that the cointegration test based on (2) provides useful additional information over the standard test procedure of Johansen (1991). Note that if ε_t is actually Student, t distributed and if there are no aberrant data structures, the estimator based on (2), of course, becomes the maximum likelihood estimator with its well-known optimality properties.

The cointegration test based on the Student t pseudo-likelihood is constructed as follows. Let $\tilde{\theta}$ and $\hat{\theta}$ denote the parameter vectors that maximize (2) with respect to θ under the null and alternative hypothesis, respectively. The cointegration test of H_r versus H_k is then given by

$$2\ln(\mathcal{L}(\hat{\theta})/\mathcal{L}(\tilde{\theta})). \tag{3}$$

As (3) is based on a ratio of two pseudo-likelihoods, we call it a pseudo-likelihood ratio (PLR) test. Note that the test of Johansen is contained as a special case if $\nu \to \infty$. The limiting distribution of this statistic is derived in Lucas (1996, 1997). For $\nu \to \infty$ the distribution collapses to the one derived by Johansen (1988, 1991).

In order to perform inference with the Student t PLR test, critical values are needed. Lucas (1996) contains a method for simulating critical values for LM-type tests, while Lucas (1997) contains a table for our LR test in case the regression model (2) contains no constant. As we want to allow for a constant in the regression model and for non-zero drift terms in the data generating process, we present a new table with critical values in this paper. As in Johansen (1991, Theorem 2.1), the critical values depend on whether the constant μ in (3) lies in the cointegrating space or not, i.e., $\alpha_\perp'\mu = 0$ or not, with $\alpha_\perp$ the orthogonal complement of α. For practical purposes, we display critical values for several values of ν. The values are presented in Table I. The columns under the heading "with drift" contain the critical values for the case $\alpha'_{\perp}\mu \neq 0$. Similarly the critical values for the case $\alpha'_{\perp}\mu = 0$ are given under heading "without drift".

< INSERT TABLE I AROUND HERE >

As an alternative to using the critical values in Table I for inference, one could use the bootstap in order to simulate the pvalues of the PLR test. Some unreported simulation experiments, however, reveal that this approach is still too time consuming for useful practical purposes.

2.2. *The construction of weights*

A useful byproduct of robust estimators is that weights are obtained for the individual observations. This can be illustrated by the simple location model

$$y_t = \mu + \varepsilon_t, \tag{4}$$

where $\varepsilon_t \overset{i.i.d.}{\sim} (0,1)$. The Student t pseudo-maximum likelihood (PML) estimator in this case solves

$$\sum_{t=1}^{T} \frac{(\nu+1)}{\nu} \cdot \frac{(y_t - \mu)}{1 + (y_t - \mu)^2/\nu} = 0, \tag{5}$$

with respect to μ. Let $\hat{\mu}$ denote the final estimate, then $\hat{\mu}$ can be interpreted as the arithmetic mean of the reweighted sample $w_t^2 y_t$, with

$$w_t^2 = (1 + (y_t - \hat{\mu})^2/\nu)^{-1} \left(T^{-1} \sum_{t=1}^{T} (1 + (y_t - \hat{\mu})^2/\nu)^{-1}\right)^{-1}. \tag{6}$$

This follows from the fact that $\hat{\mu} = T^{-1} \sum_{t=1}^{T} w_t^2 y_t$ is the same as $\sum_{t=1}^{T}(w_t^2 y_t - w_t^2 \hat{\mu}) = 0$, which is easily seen to satisfy (5). Note that w_t^2 is not bounded from above by 1, but rather by $(\nu + 1)/\nu$. Also note that $\hat{\mu}$ can be interpreted as the OLS estimator of μ in the weighted regression model

$$w_t y_t = w_t \mu + w_t \varepsilon_t. \tag{7}$$

One can interpret w_t as the weight for the observation at time t. A low value of w_t indicates that the observation does not correspond to the general pattern in the data. Alternatively, one can interpret w_t as the inverse of the standard deviation of the error term. In that case, multiplying the observations by w_t and computing the OLS estimator as in (7) can be seen as a generalized least-squares (GLS) correction for the possible presence of heteroskedasticity. In this GLS interpretation, a large value of w_t^{-1} indicates that the error term at time t has a high variance and, therefore, has to be downweighted. These two interpretations can be usefully exploited in practical occasions.

Similar to the Student t PML estimator for (4), the Student t PML estimator for (1) can be regarded as the Gaussian PML estimator for a weighted version of model (1), with the weights given by

$$w_t = \left(\frac{\nu + k}{\nu + \varepsilon_t' V^{-1} \varepsilon_t} \right)^{1/2}. \tag{8}$$

To see this, note that the first order condition defining the PML estimator is given by,

$$\sum_{t=1}^{T} \frac{(\nu + k)\varepsilon_t(\theta)'}{\nu + \varepsilon_t(\theta)' V^{-1} \varepsilon_t(\theta)} \cdot \frac{\partial \varepsilon_t(\theta)}{\partial \theta'} = \sum_{t=1}^{T} w_t^2 \varepsilon_t(\theta)' \frac{\partial \varepsilon_t(\theta)}{\partial \theta'}. \tag{9}$$

where $\varepsilon_t(\theta) = \Delta y_t - \alpha\beta' y_{t-1} - \Phi_1 \Delta y_{t-1} - \ldots - \Phi_{p-1} \Delta y_{t-p+1} - \mu$. In obtaining (9), we have assumed for simplicity that V is known. If $\nu \to \infty$, $w_t \equiv 1$ and (9) reduces to the first order condition defining the Gaussian PML estimator of Johansen (1991). Consequently, the first order condition for the Student t PML estimator can be interpreted as an observation weighted version of the first order condition of the Gaussian PML estimator.

Equations (8) and (9) clearly illustrate that the Student t PML estimator satisfies the two objectives formulated in Section 1. Observations with unusually large values of ε_t automatically receive a smaller weight. Furthermore, these weights are available for inspection after the estimates and test statistics have been computed. Note that the weighting scheme implied by (8) operates in the intuitively correct way if the aberrant observation happens to give a clear signal about the underlying error correction mechanism. To see this, we assume a single large value for ε_t at time t_0. This large innovation induces a large equilibrium error $\beta' y_{t0}$, which, in turn, causes a highly leveraged observation in (2) at time $t_0 + 1$. If ε_t contains no further abnormal values for $t > t_0$, (8) shows that only one regression observation receives a smaller weight, as there is only one outlying ε_t. The information on the error correction mechanism contained in the subsequent observations is fully exploited. By contrast, the Gaussian PML estimator also exploits the information contained in these latter observations, but it fails to correct for the outlying value of ε_t initiating the mechanism.

In order to decide whether an observation of a multivariate process receives an extraordinarily small weight, we employ the following methodology. Under the assumption that the ε_t are standard normally distributed, $\varepsilon_t' V^{-1} \varepsilon_t$ has a χ^2 distribution with k degrees of freedom. We choose to assume that ε_t is normally distributed, since this corresponds to the situation in which there are no outliers or structural breaks. Needless to say that other choices for the distribution of ε_t lead to other decision rules. Let ~(0.005) denote the one half per cent critical value for the χ^2 distribution with k degrees of freedom. Weights are then found to be extraordinarily small if $w_t^2 \leq (\nu + k)/(\nu + c_k$ (0.005)). For example, for $\nu = 5$ and $k = 2$, this means that observations with weights smaller than approximately 0.67 deserve a closer inspection. Of course, other quantiles of the χ^2 distribution lead again to other decision rules. In the illustration and the application in the following sections we evaluate the weights using the estimates under the alternative H_k. Unreported experiments reveal that one can also use the estimates under the null hypothesis H_r in most practical circumstances. Using the estimates under the alternative, however, precludes a bias in the weights due to a misspecified cointegrating rank r.

3. SIMULATED DATA

In this section we consider the implications of using robust estimation methods in practice. We do this by means of simulated data. First, we describe the different patterns in the weights that result from applying the robust estimation in several settings of practical interest. Second, we describe the sensitivity of the cointegration and outlier testing procedures to the choice of the tuning constant ν.

3.1. Several simulated illustrative examples

In order to illustrate the effect of different types of data irregularities on the patterns of the weights, we discuss four examples. The first example concerns an additive outlier, the second considers a temporary level shift, the third example deals with a variance shift, and the fourth example illustrates the effect of a patch of innovative outliers. All examples consider bivariate time series for 100 observations. These time series are generated from the model

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{pmatrix} 3.7 \\ -0.3 \end{pmatrix} (-0.060,\ \ 0.075,\ \ 0.018) \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ 1 \end{pmatrix} + \varepsilon_t, \tag{10}$$

where ε_t is normally distributed with mean zero and covariance matrix

$$\begin{pmatrix} 1.00 & 0.28 \\ 0.28 & 0.20 \end{pmatrix}.$$

The parameter values in (10) are derived from the estimated parameters in a VAR model for long-run and short-run interest rates, see Franses and Lucas (1997).

The results of our small simulation experiment for the additive outlier (AO) case are presented in the two graphs in the first row of Figure 1. Our method is as follows. First, a realization of (10) is drawn. Next, we replace $y_{2,49}$ by $y_{2,49} - 2$. This creates a series with an additive outlier at $t = 49$. The two series y_{1t} and y_{2t} are presented in the upper-left graph of Figure 1. For these two series, an unrestricted VAR(1) model with constant term is estimated using the Student t PML estimator with $\nu = 5$. The parameter estimates of this model are used to construct estimates of the weights in (8). Note that these weights apply to both time series. The weights, along with a critical value of 0.67, are presented in the upper-right panel of Figure 1. It appears that four observations obtain a weight below the critical threshold. These observations are marked by a circle in the upper-left panel of Figure 1. Clearly, for observations 49 and 50 the small weights seem to be due to the additive outlier in the y_{2t} series. Note that the additive outlier can in this case be viewed as a large, negative innovative outlier in period 49, followed by a large, positive innovative outlier in period 50, because we consider a VAR(l) model. This generalizes to VAR models of order p, in which case a patch of $p + 1$ low weights can be expected in case of an AO. By contrast, for innovative outliers one expects that only one observation receives a smaller weight, see the explanation in Subsection 2.2. Notice that the patch of outlying (p + l)-tuples $(y_t, \ldots, y_{t-p})$ for one AO may suggest a higher order for the VAR model in practice. Hence, we may overestimate the proper order of the VAR model. In this paper we abstain from a discussion of model order selection in the presence of aberrant observations. Therefore, in our empirical section below we evaluate the properties of VAR models of various orders.

It is also interesting to look at the effect of the additive outlier on the PLR test and on the estimates of α and β, see (1). For the series without the additive outlier, the Gaussian PLR test for $r \leq 1$ equals 1.57, while the Student t based PLR test is 1.86. If we add the additive outlier, the Gaussian PLR test is increased by about 40%, while the Student t PLR test is inflated by only about 7.5%. This illustrates the sensitivity of the Gaussian based procedure and the insensitivity of the Student t approach to isolated aberrant observations.

< INSERT FIGURE 1 AROUND HERE >

The second example considers a temporary level shift. We use the same series as in the case of the additive outlier. We restore the original series by removing the AO at t = 49, and replace y_{2t} by y_{2t} + 3 for the observations at t = 45,. . . ,53. The results are presented in the bottom two graphs of Figure 1. Looking at the dashed line, the level shift is clearly visible. The weights reveal that four observations require a closer inspection. Similar to the first example, the observations at t = 18 and t = 71 seem not very relevant. The temporary level shift is clearly indicated by the two small weights for period 45 and 54. In period 45, the y_{2t} series has jumped upwards, causing the observation to receive a small weight. Similarly, in period 54 the series has jumped back to its original level, resulting in a small weight at t = 54. A temporary level shift is thus only indicated by the small weights at the starting date and at the ending date of the shift. Of course, one may now proceed with an application of the techniques developed in Tsay (1988) to formally distinguish between an AO and a level shift model, but this is not pursued here.

Again we also consider the effect of the temporary level shift on the cointegration testing procedure. Without the level shift, the Gaussian test gives the values 17.31 and 1.57 for the null hypotheses $H_0 : r \leq 0$ and $H_1 : r \leq 1$, respectively. The Student t test gives the values 17.02 and 1.86, respectively. If we add the level shift, the Gaussian test values increase by 17% and 62% to 20.18

and 2.55, respectively. The robust test, by contrast, gives the values 17.14 and 0.34. Both these numbers show that the Student t based test is much less biased to the alternative hypothesis of stationarity than the Gaussian test.

For the third example of a variance shift, a new realization of (10) is drawn with the modification that from t = 79 onwards, ε_{2t} is replaced by $3\varepsilon_{2t}$. The results are presented in the top graphs of Figure 2. It is clear that at the end of the sample there are several observations that are not described adequately by the model. The fact that not all observations from period 79 onwards receive a small weight is due to the fact that ε_{2t}, even when multiplied by a factor three, can still be quite small. Diagnostic tests for heteroskedasticity may now be useful in a next step of the empirical analysis.

< INSERT FIGURE 2 AROUND HERE >

The fourth and final example considers a patch of innovative outliers. Again, a new realization of (10) is drawn with the modification that for $t = 79, \ldots, 99$, ε_{2t} is replaced by $\varepsilon_{2t} + \delta_t$, with $\{\delta_t\}_{79}^{99}$ a set of independently and identically distributed random variables with $P(\delta_t = -3) = P(\delta_t = 3) = 0.5$. The results are presented in the bottom graphs of Figure 2. The patch of innovative outliers clearly emerges from the cloud of circles in the right corner of the lower-left panel of Figure 2, as well as from the large number of small weights for the last observations of the sample (see the lower-right panel).

Both the variance shift and the patchy innovative outliers result in aberrant observations that satisfy the dynamics of the model. Therefore, these types of aberrant data structures do not result in a large bias for the PLR test, either based on the Gaussian or on the Student t distribution. If the increase in variance or the magnitude of the outliers is large enough, it is clear that power can be gained by concentrating on the low-variance part of the sample. As can be seen from the weights in Figure 2, this is precisely what the Student t PLR test does.

To summarize the four illustrative examples in this subsection, we conclude that the Student t test is less biased towards stationarity than the Gaussian PLR test. Moreover, a plot of the weights implied by the robust estimation method may provide very useful information about which observations do not seem to fit into the model. Closer inspection of these observations is then needed, e.g., using the methods in Tsay (1988), and this might result in a respecification of the model or in a re-interpretation of cointegration results. In some circumstances, the weights can even be used to identify the type of model failure, as in the relatively simple case of an additive outlier. In other cases, this is much more difficult, as can be seen from the graphs for variance shifts and patches of innovative outliers. Balke (1993) shows that with formal tests one can also find difficulties in distinguishing between the various outlier types.

3.2. *The choice of* ν

Our robust cointegration approach outlined in Section 2 involves two practically relevant choices for key parameters. The first is the value of the tuning constant ν, which was set to 5 in the illustrative examples in Subsection 3.1. The second parameter is the critical value of the weights w_t. In this subsection we evaluate the sensitivity of the results from our approach to various choices of ν. As a side result, we establish the sensitivity of the standard Johansen procedure to aberrant observations. We design the following simulation experiment. We generate 1,000 replications of 100 observations from the data generating processes (DGPs)

$$
\begin{aligned}
&(i): \quad \begin{cases} \Delta y_{1t} = \varepsilon_{1t} \\ \Delta y_{2t} = \varepsilon_{2t} \end{cases}, \\
&(ii): \quad \begin{cases} \Delta y_{1t} = \varepsilon_{1t} \\ \Delta y_{2t} = -0.2(y_{2,t-1} \quad y_{1,t-1}) + \varepsilon_{2t} \end{cases}, \qquad (11) \\
&(iii): \quad \begin{cases} \Delta y_{1t} = -0.2 y_{1,t-1} + \varepsilon_{1t} \\ \Delta y_{2t} = -0.2 y_{2,t-1} + \varepsilon_{2t} \end{cases},
\end{aligned}
$$

where the VAR(1) regression model has two unit roots in case (i) ($r = 0$), one unit root and one cointegrating relation in case (ii) ($r = 1$), and no unit roots (stationarity, $r = 2$) in case (iii). We contaminate these data with a certain percentage of additive outliers of a specific size and we test for cointegration in a VAR(l) model with an unrestricted constant term using our PLR test with various values of ν.

< INSERT TABLE II AROUND HERE >

We present the simulation results and more details of our experiment in Table II. As expected, the power of the cointegration test is low, since the stationary relations in our DGPs have a root of 0.8, see (iii) in (11). The rejection frequencies in Table II can thus be decomposed in two parts. The first part is due to the true power of the PLR test. This part is given in the row $\zeta = 0$, which is the case without outliers. The second part is due to the bias caused by the additive outliers. This is given by the difference between the rows for $\zeta \neq 0$ and $\zeta = 0$.

Table II clearly illustrates the price one has to pay for using the outlier robust method. If there are no outliers ($\zeta = 0$) the G aussian method selects the correct model more often than the Student t procedure. This is evident, as the Gaussian method is optimal in that case. The loss in power does not seem dramatic for all values of ν considered.

If we add outliers to the samples, we note that the size of the Gaussian PLR test is severely distorted. The test is clearly biased towards stationarity. For example, for the bivariate random walk case ($r = 0$), the actual size of the Gaussian test ($\nu = \infty$) increases from the nominal level of 5% for $\zeta = 0$ to 15% and 58% for 1% and 5% of additive outliers ($\zeta = 7$), respectively (sum the entries for $\zeta = 7$, $r = 1$, $\nu = \infty$, and? $= 1, 2)$. Similarly, in the case of one cointegrating vector $(r = 1)$, the size increases to 10% and 36% for 1% and 5% of additive outliers, respectively (see the entry for $\zeta = 7$, $r = 1$, $\hat{r} = 2$, and $\nu =$ co). In the stationary case, the bias towards stationarity leads to the correct inference on the cointegrating rank of the system. The effect of outliers manifests itself in this case through the lower degree of persistence of the shocks in the system. Using (iii) from (11), one can easily show that the roots of the system converge in probability to $0.8/(1 + 0.018\zeta^2)$ if the fraction of outliers is 5%. This leads to biases of 14%, 32%, and 47% for $\zeta = 3, 5, 7$, respectively. In small samples as the ones used to construct Table II, the biases with respect to the no-outlier case $\zeta = 0$ reduce to 8%, 19%, and 30%, respectively, which is still considerable. These findings corroborate the results from the literature, stating that in the presence of additive outliers the Gaussian-based procedure of Johansen (1988, 1991) tends to find models that are "too stationary," see Franses and Haldrup (1994) and Hoek et al. (1995).

Except for the stationary model, the size distortions of the Student t test are less than those of the Gaussian test. The effect becomes more pronounced if the outliers increase in magnitude. So our robust method tends to find less spurious cointegration. The results further show that the size distortion is increasing in ν. This is intuitively clear, as the Student t distribution starts to resemble the Gaussian distribution for larger values of ν. Stated differently, the weights in (8) are monotonically increasing in ν for large values of ε_t, such that for higher ν the observation weights of the Student t procedure approach the unit observation weights of the Gaussian estimator. For DGP (iii), the robust approach evidently results in less stationary models than the Gaussian approach. As explained earlier, however, the correct inference resulting from the Gaussian test is more a result of the bias in the test than of the power of the test. This is clearly seen if we look at the roots of the system. For $\nu = 5$ and 5% additive outliers, the simulations used to construct Table II indicate biases with respect to the no-outlier case $\zeta = 0$ of 6%, 12%, and 17% for $\zeta = 3, 5, 7$, respectively. These biases are considerably smaller than the corresponding biases of the Gaussian estimator. We conclude that the Student t PLR test provides at least some protection against the distortionary effect of additive outliers.

The extent to which the choice of ν influences the results in Table II seems limited. Obviously, ν should not be chosen too high if one wants protection to outliers. Choices in the range $\nu = 3, \ldots, 7$ all seem acceptable, however. Lower values of ν provide somewhat more protection against aberrant observations, but at the cost of a somewhat higher loss in power if there are no outliers ($\zeta = 0$). It should be mentioned here that (unreported) simulation results for other contamination percentages and for a VAR(2) model instead of a VAR(l) model yield qualitatively similar results, i.e., our method does not seem particularly sensitive to the choice for a particular value of $\nu < 10$.

< INSERT TABLES III AND IV AROUND HERE >

Similar conclusions can be drawn for the second important step, i.e., the decision on which observations have weights that differ significantly from one. The second part of our above simulation experiment involves an investigation whether our method is able to detect x per cent influential observations when there are indeed x per cent of such data points.

The relevant results are summarized in Tables III and IV for 1% and 5% AOs, respectively. The entries in these tables are fairly nonstandard and need some explanation. For a given DGP and given values of ζ and ν, we have the entry

average percentage of observations **correctly** classified as outlier	average percentage of observations **incorrectly** classified as 'clean'
average percentage of observations **incorrectly** classified as outlier	average percentage of observations **correctly** classified as 'clean'

So we would like the off diagonal elements of these entries to be as small as possible. First note that we expect the average number of outliers in our samples to be 200 . p_c, with p_c = 1%, 5% the fraction of additive outliers. This was explained in the previous subsection. Obviously, the weights of the Gaussian estimator do not detect the aberrant observations, as the weights are identically equal to unity. Furthermore, as expected, the robust method does a much better job at detecting the irregular data patterns. The method seems to work better when the size ζ of the AOs is larger. It is important to note that the results in Tables III and IV again indicate that our empirical findings based on our outlier robust method are, in turn, robust to the choice of ν. The number of correct classifications made by the Student $t(\nu = 5)$ estimator, however, seems to be somewhat higher than for the other choices of ν.

In sum, the simulation results in this section suggest that for many practical purposes, we may consider setting ν equal to 5. This gives a test procedure that provides some protection against the adverse effects of aberrant data structures. Moreover, the method produces useful diagnostics in the form of observation weights in order to assess which observations are not described by the model.

4. EMPIRICAL ILLUSTRATION

In this section we illustrate the practical usefulness of our outlier robust cointegration analysis by analyzing annual observations of the US and Finland Consumer Price Indices and the US/Finland nominal exchange rate, 1900-1988. All data are transformed to logs. Perron and Vogelsang (1992), Franses and Haldrup (1994), and Hoek et al. (1995) also consider these data and in particular the real exchange rate between Finland and the US. This latter series is depicted in the left panel of Figure 3. The interesting economic question is whether the real exchange rate is stationary or not. We perform two analyses: one with the Gaussian based testing procedure of Johansen ($\nu = \infty$), and one with the Student t based likelihood with five degrees of freedom ($\nu = 5$). We compare the empirical results obtained with the different methods. If the two methods give different results, closer inspection of the data and the model can be warranted.

< INSERT FIGURE 3 AROUND HERE >

Consider the test results in Table V. The Akaike and the Schwarz criterion suggest the adequacy of a VAR(2) model for both the robust and nonrobust method. We also report the results for other VAR orders for completeness, because it is yet unknown how model selection is affected by outliers and by the use of robust estimation techniques. For the VAR(2), the nonrobust method indicates that there is one cointegrating relation. This relation can approximately be identified as the real exchange rate (i.e., cointegrating vector (1, -1, 1)), since the estimated cointegrating vector is $(1, -0.87, 0.82)$. The robust method for $\nu = 5$, by contrast, indicates no significant cointegrating relationships for most values of p, including the selected value $p = 2$. It is noticeable, however, that in this robust case the estimated cointegration vector is $(1, -0.88, 0.85)$, which is again close to the real exchange rate.

< INSERT TABLE V AROUND HERE >

In order to investigate the cause of the difference between the results of the Gaussian and the Student $t(\nu = 5)$ based cointegration tests, the implied weights of the robust estimator for $\nu = 5$ are plotted in the right panel of Figure 3. With a 0.5% critical value corresponding to a normal distribution of the underlying errors, we assign significantly less weight to the observations in 1915-1921, 1932, 1945, 1946, 1949, and 1957. During 1915-1925, we have the events of World War I, the Finnish independence (1917), and the Finish civil war (1918). This clearly accounts for the first set of aberrant observations, during which Finland experienced a floating exchange rate regime. The small weight for 1932 marks the crisis of the thirties and also indicates the middle of a relatively short floating exchange rate regime (1931-1933) between the period of the Gold standard (1926–1931) and the Pound standard (1933-1939). The small weights for 1945 and 1946 probably reflect the impact of World War II. The year 1949 and 1957 both fall within the Bretton-Woods system. At the start, Finland experienced two large devaluations of 18% and 44% in July and September 1949, respectively. In September 1957, there was another large devaluation of 39%. Strangely enough, the relatively large devaluation of 31% in October 1967 does not show up in the results. This holds both for our robust detection method and for the detection method of Cheng and Liu (1993) as employed in Franses and Haldrup (1994).

It should perhaps be mentioned that we also experimented with different critical values for the weights. For example, if the error process is assumed to be Student *t(5)* instead of Gaussian, we obtain significant weights for 1917-1921, 1932, 1945, 1946, 1949, and 1957. Hence, our decision on which data are aberrant does not seem to be influenced much by our cut-off measure. Finally, notice that our finding that the US/Finland real exchange rate is nonstationary once we have taken care of outliers, clearly matches the findings in Franses and Haldrup (1994) and Hoek et al. (1995), where a less sophisticated route is followed.

In sum, our analysis illustrates that outlier robust procedures provide useful additional information for validating the results obtained with a traditional non-robust analysis. In our empirical example, cointegration found between consumer price indices of Finland and the US and the nominal exchange rate between these countries, hinges on the presence of only a few outliers.

5. CONCLUDING REMARKS

Standard unit root and cointegration tests are sensitive to outliers and structural changes in the data. Additionally, one may expect tests for stationarity, such as those proposed in, e.g., Leybourne and McCabe (1994), to suffer from the same problems. Although it is true that such standard tests are not designed for nonstandard cases, it is our experience that aberrant data frequently occur in practice, and hence that our outlier robust procedures can be used to validate the results of a traditional analysis. Using both synthetic and empirical data, we showed that our outlier robust method is more resistant to aberrant data structures than the Gaussian based method of Johansen (1991). Moreover, the robust approach produces a valuable diagnostic tool in the form of observation weights. A graph of these weights provides useful additional information and we recommend its use in practice. Knowledge about which observations are not well described by the model clearly helps the practitioner in choosing alternative model specifications and/or including additional variables.

One can argue that the outlier robust method is approximately equal to traditional Gaussian based (OLS) analysis with some dummy variables. The implementation of outlier detection methods as in Tsay (1988) in a multivariate context with possible integrated series, however, is far from trivial. As yet, there seems to be no practical way of taking the decisions on the type, magnitude, and dynamic patterns of various outliers simultaneously in this multivariate context. Compared to the traditional approach augmented with dummies, the outlier robust method has two main advantages. First, the outlier robust procedure automatically incorporates the identification and handling of different types of outliers into the estimation stage. Therefore, the applied econometrician does not have to run many regressions with many different variations of dummies in order to check the sensitivity of the obtained results. Moreover, the complications in limiting distributions of unit root tests arising from the fact that dummies are constructed on the basis of in-sample information (see Christiano (1992)) are avoided. The second advantage of the outlier robust procedure over the OLS with dummies approach is that additional information can be obtained by inspecting the observation weights produced by the robust estimator. Both these advantages seem to advocate the use of outlier robust unit root and cointegration tests in empirical studies.

ACKNOWLEDGEMENTS

We thank Pekka Pere and Carolina Sierimo for kindly giving us the necessary details on Finnish (monetary) history. The first author thanks the Royal Netherlands Academy of Arts and Sciences for its financial support. Both authors thank an Associate Editor and two anonymous referees for their very extensive and helpful comments and suggestions for improvement.

Econometric Institute, Erasmus University Rotterdam, P. 0. Box 1738, 3000DR Rotterdam, The Netherlands

and

Financial Sector Management, Free University Amsterdam, De Boelelaan 1105, 1081HV, Amsterdam, The Netherlands

REFERENCES

Balke, N.S. (1993): "Detecting level shifts in time series," *Journal of Business and Economic Statistics* 11, 81-92.

Chen, C., and L.M. Liu (1993): "Joint estimation of model parameters and outlier effects in time series," *Journal of the American Statistical Association 88, 284-297.*

Christiano, L.J. (1992): "Searching for a break in GNP," *Journal of Business and Economic Statistics* 10, 237–250.

Franses, P.H., and N. Haldrup (1994): "The effects of additive outliers on tests for unit roots and cointegration," *Journal of Business* and *Economic Statistics* 12, 471-478.

Franses, P.H., and A. Lucas (1997): "Outlier robust cointegration analysis of Dutch interest rates," *Proceedings of the 1996 Meeting of the American Statistical Association, (to appear).*

Gouriéroux, C., A. Monfort, and A. Trognon (1984): "Pseudo maximum likelihood methods: theory," *Econometrica 52,* 681-700.

Hoek, H., A. Lucas, and H.K. van Dijk (1995): "Classical and Bayesian aspects of robust unit root inference," *Journal of Econometrics 69, 27-59.*

Johansen, S. (1988): "Statistical analysis of cointegration vectors," *Journal of Economic Dynamics and Control* 12, 231-254.

Johansen, S. (1991): "Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models," Econometrica 59, 1551-1580.

Johansen, S., and K. Juselius (1990): "Maximum likelihood estimation and inference on cointegration — with applications to the demand for money," *Oxford Bulletin of Economics and Statistics 52,* 169–210.

Leybourne, S.J., and P.M. McCabe (1994): "A consistent test for a unit root," *Journal of Business and Economic Statistics* 12, 157-166.

Lucas, A. (1995a): "An outlier robust unit root test, with an application to the extended Nelson-Plosser data," *Journal of Econometrics 66,* 153-173.

Lucas, A. (1995b): "Unit root tests based on M estimators," *Econometric Theory* 11, 331-346.

Lucas, A. (1996): "Inference on cointegrating ranks using pseudo LM tests," VU Research Memorandum 42, Free University Amsterdam.

Lucas, A. (1997): "Cointegration testing using pseudo likelihood ratio tests," *Econometric Theory* 13, 149-169.

Perron, P. (1989): "The great crash, the oil price shock, and the unit root hypothesis," Econometrica 57, 1361-1401.

Perron, P., and T.J. Vogelsang (1992): "Nonstationarity and level shifts with an application to purchasing power parity," Journal *of Business* and Economic *Statistics* 10, 301-320.

Tsay, R. (1988): "Outliers, level shifts, and variance changes in time series," *Journal of Forecasting 7, l-20.*

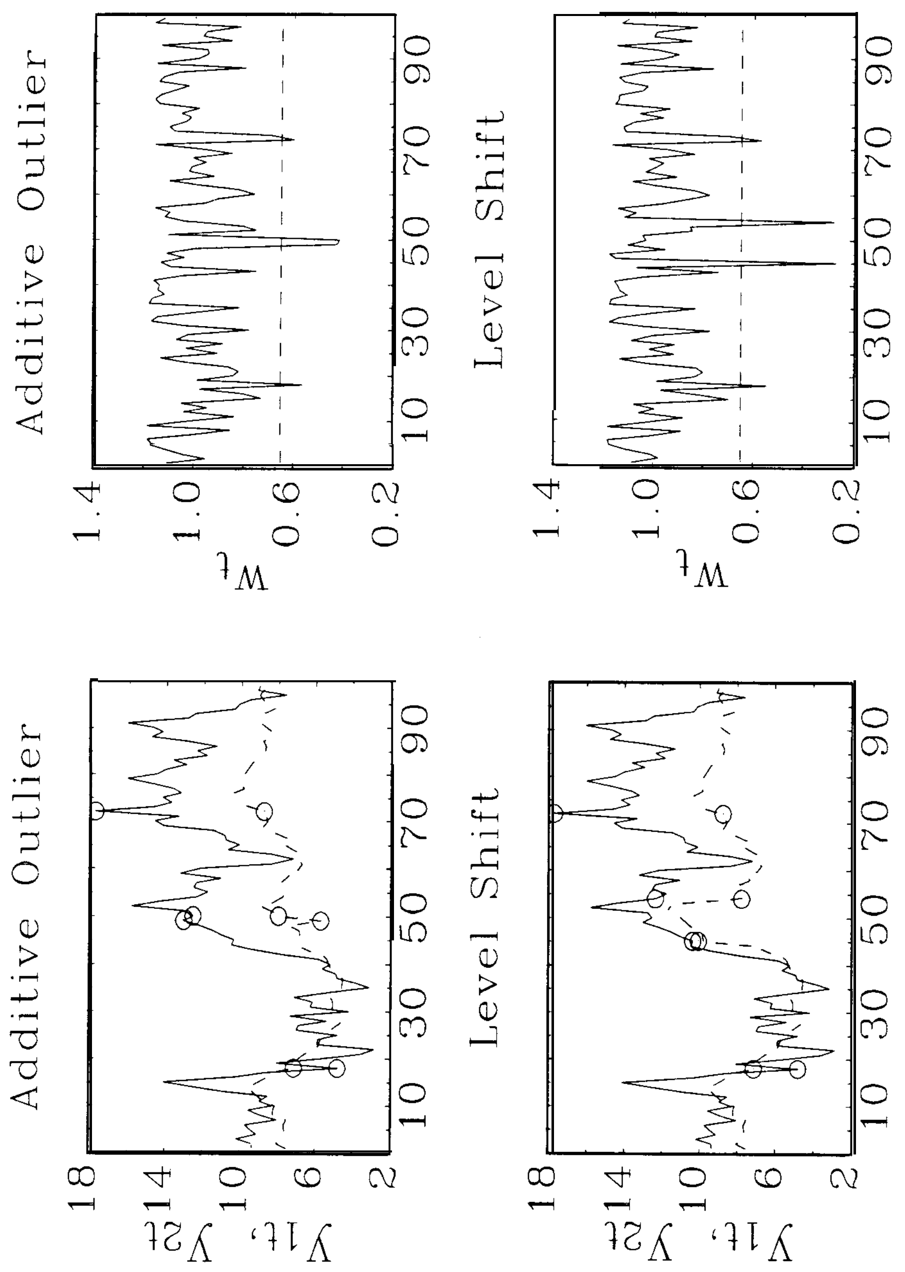

Figure 1.— Weights (w_t) of the Student t Pseudo Maximum Likelihood Estimator for an Additive Outlier and a Temporary Level Shift corresponding to a Bivariate Time Series (y_{1t}, y_{2t})

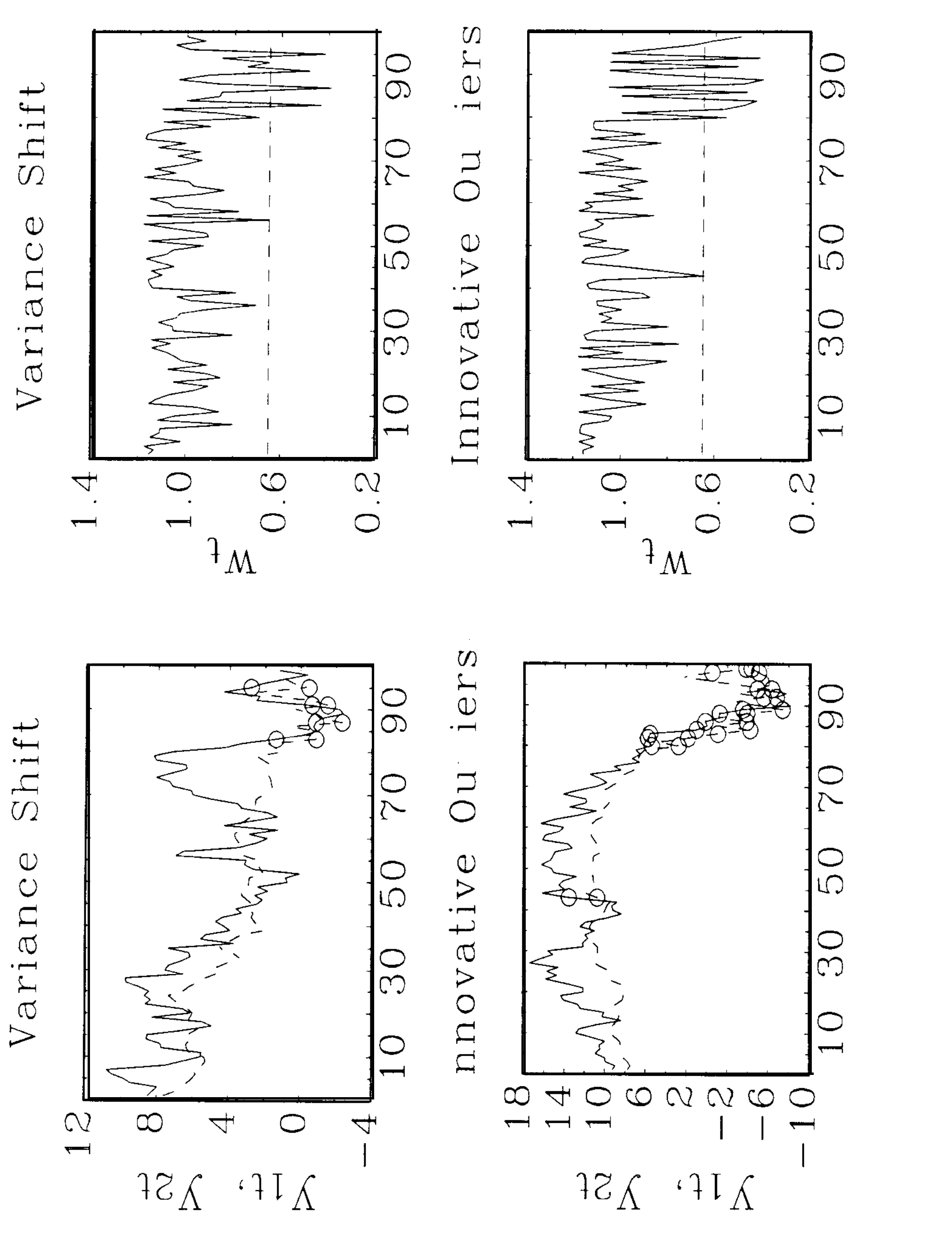

Figure 2.— Weights (w_t) of the Student t Pseudo Maximum Likelihood Estimator for a Variance Change and a Patch of Innovative Outliers corresponding to a Bivariate Time Series $(y_{1t},\ y_{2t})$

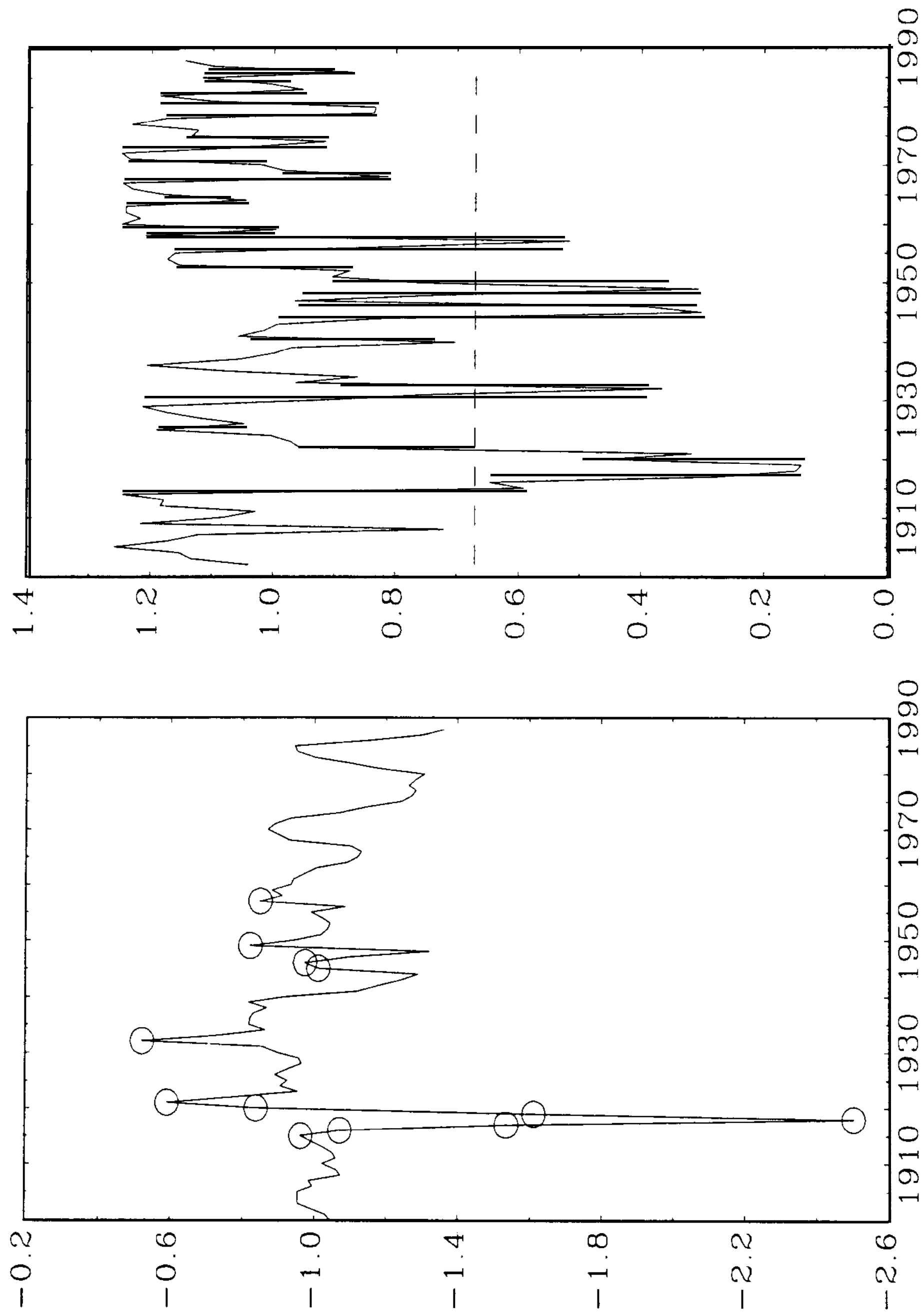

Figure 3.— Weights obtained with $\nu = 5$ and the Finland/US Real Exchange Rate (CPI Deflated)

TABLE I

QUANTILES OF THE STUDENT t BASED PSEUDO LIKELIHOOD RATIO TEST

$k-r$	ν	with drift				without drift			
		0.80	0.90	0.95	0.99	0.80	0.90	0.95	0.99
	∞	1.7	2.9	4.1	7.1	4.9	6.4	7.8	11.2
	10	1.8	3.0	4.3	6.6	5.0	6.5	8.0	11.3
1	7	1.8	3.1	4.4	7.1	5.0	6.5	8.1	11.7
	5	1.9	3.2	4.4	6.8	5.1	6.6	8.5	11.7
	3	2.0	3.4	5.0	7.5	5.1	7.1	9.0	12.1
	∞	11.5	13.6	15.7	20.6	13.6	15.6	18.2	21.8
	10	11.9	14.1	16.3	22.1	13.8	16.6	18.6	23.1
2	7	11.8	14.3	16.5	23.2	14.0	16.9	19.1	23.7
	5	12.1	14.6	17.2	23.5	14.2	17.2	19.8	24.1
	3	12.8	15.6	18.3	24.3	15.0	18.0	20.6	25.7
	∞	23.9	27.3	30.1	35.5	26.1	29.2	32.6	37.3
	10	24.9	28.1	30.7	36.9	26.4	29.9	33.1	38.4
3	7	25.2	28.8	31.2	37.6	26.8	30.3	33.7	39.6
	5	25.5	29.6	31.7	38.6	27.5	30.9	34.3	41.0
	3	26.5	30.7	33.5	40.8	28.6	32.2	35.9	43.2

The table contains the quantiles of the PLR test based on the Student t pseudo likelihood. ν denotes the degrees of freedom parameter of the pseudo likelihood. The entries are based on 1,000 Monte-Carlo simulations. In each simulation, a $k-r$ dimensional Gaussian random walk is generated, possibly with nonzero drift. Next, the PLR test of the hypothesis of zero versus $k-r$ cointegrating relations was computed, using model (1). The constant in (1) enters unrestrictedly in the regression model. The left set of quantiles is based on a data generating proces with a drift term equal to the standard deviation of the innovations. The second set of quantiles is based on a data generating process with zero drift. Therefore, the quantiles in the present table can be compared to the quantiles in Tables Al and A2 of Johansen and Juselius (1990), respectively.

TABLE II

REJECTION FREQUENCIES OF THE PLR TEST WITH ADDITIVE OUTLIERS

r	ζ	$\hat{r}$														
		$\nu = \infty$			$\nu = 10$			$\nu = 7$			$\nu = 5$			$\nu = 3$		
		0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
		1% additive outliers														
0	0	95	4	1	95	5	1	95	5	1	95	5	1	95	5	0
0	3	94	5	1	94	5	1	95	5	1	94	5	1	94	5	1
0	5	91	8	1	93	6	1	94	6	1	94	5	1	94	5	1
0	7	85	13	2	92	7	1	94	6	1	94	6	1	94	5	1
1	0	58	37	5	59	36	5	61	34	5	61	34	4	64	33	3
1	3	50	44	6	54	40	6	56	38	6	57	38	5	59	37	4
1	5	40	54	7	49	44	7	53	41	6	54	40	6	58	37	5
1	7	31	58	10	46	47	7	50	43	7	53	42	6	58	38	5
2	0	6	21	72	8	23	69	9	25	65	11	29	60	15	35	50
2	3	4	18	78	6	22	73	7	23	70	9	26	65	13	32	55
2	5	3	16	81	5	21	75	6	22	72	8	25	67	12	31	57
2	7	3	15	82	4	20	76	6	22	72	7	25	67	11	32	57
		5% additive outliers														
0	0	96	4	0	96	4	0	96	4	0	96	4	0	96	4	0
0	3	88	11	1	90	9	1	91	8	1	91	9	1	92	8	0
0	5	67	29	4	80	18	2	83	15	2	86	13	1	89	10	1
0	7	42	47	11	72	25	4	78	19	3	83	15	2	88	11	1
1	0	60	36	4	61	36	3	63	33	3	64	33	3	67	30	3
1	3	30	61	9	38	53	8	41	52	7	44	50	7	48	46	6
1	5	10	70	20	20	66	14	26	63	11	32	58	10	42	51	7
1	7	3	61	36	14	67	20	20	65	15	26	62	12	39	53	8
2	0	7	24	69	8	26	66	10	28	62	11	30	59	15	36	48
2	3	1	9	90	1	13	86	2	16	82	3	20	78	5	28	67
2	5	0	4	96	1	8	92	1	10	89	1	14	84	3	22	74
2	7	0	3	97	0	5	94	0	8	92	1	11	88	3	21	76

NOTE: The table contains the percentage of times in 1,000 Monte-Carlo replications that a cointegrating rank of $\hat{r}$ is found based on the Student t density. ν denotes the degrees of freedom parameter in the Student t pseudo-likelihood. r denotes the true cointegrating rank. For $r = 1$, the roots of the system are 1 and 0.8, respectively. For $r = 2$, both roots are 0.8. The original series are contaminated in the following way. For given y_t, y_t is replaced by $y_t + \zeta u_t$ with probability p, where u_t denotes a drawing from an i.i.d. process that is uniformly distributed over the unit circle. Otherwise, y_t is left unchanged. The table contains results for $p = 0.01$ and $p = 0.05$, respectively.

TABLE III

OUTLIER DETECTION USING THE STUDENT t WEIGHTS UNDER 1% ADDITIVE OUTLIER CONTAMINATION

r	ζ	ν = ∞	10	7	5	3
3		0 | 2 / 0 | 98	1 | 1 / 1 | 97	1 | 1 / 1 | 97	1 | 1 / 1 | 97	1 | 1 / 2 | 96
5		0 | 2 / 0 | 98	2 | 0 / 1 | 97	2 | 0 / 1 | 97	2 | 0 / 1 | 97	2 | 0 / 2 | 96
7		0 | 2 / 0 | 98	2 | 0 / 1 | 97	2 | 0 / 1 | 97	2 | 0 / 1 | 97	2 | 0 / 2 | 96
3		0 | 2 / 0 | 98	1 | 1 / 1 | 97	1 | 1 / 1 | 97	1 | 1 / 2 | 96	1 | 1 / 3 | 95
5		0 | 2 / 0 | 98	2 | 0 / 1 | 97	2 | 0 / 1 | 97	2 | 0 / 2 | 96	2 | 0 / 3 | 95
7		0 | 2 / 0 | 98	2 | 0 / 1 | 97	2 | 0 / 1 | 97	2 | 0 / 2 | 96	2 | 0 / 3 | 95
3		0 | 2 / 0 | 98	1 | 1 / 1 | 97	1 | 1 / 1 | 97	1 | 1 / 2 | 96	1 | 1 / 3 | 95
5		0 | 2 / 0 | 98	1 | 1 / 1 | 97	2 | 0 / 1 | 97	2 | 0 / 2 | 96	2 | 0 / 3 | 95
7		0 | 2 / 0 | 98	2 | 0 / 1 | 97	2 | 0 / 1 | 97	2 | 0 / 2 | 96	2 | 0 / 3 | 95

NOTE: The table contains the average (over the Monte-Carlo replications) of the percentage of outliers that are (not) detected. A 'cell' in the table consists of four numbers, indicating

average percentage of observations **correctly** classified as outlier	average percentage of observations **incorrectly** classified as 'clean'
average percentage of observations **incorrectly** classified as outlier	average percentage of observations correctly classified as 'clean'

The simulation setup is described in the note to Table II. The percentage of additive outliers is 1.

TABLE IV

OUTLIER DETECTION USING THE STUDENT t WEIGHTS UNDER 5% ADDITIVE OUTLIER CONTAMINATION

r	ζ	$\nu = \infty$		10		7		5		3	
3		0	10	3	7	3	7	4	5	5	5
		0	90	0	90	0	90	1	90	1	89
5		0	10	7	3	8	2	8	2	9	1
		0	90	0	90	0	90	0	90	1	89
7		0	10	9	1	9	1	10	0	10	0
		0	90	0	90	0	90	0	90	1	89
3		0	10	3	7	3	7	4	6	5	5
		0	90	0	90	0	90	1	89	1	89
5		0	10	6	4	7	3	8	2	8	2
		0	90	0	90	0	90	0	90	1	89
7		0	10	8	2	9	1	9	1	9	1
		0	90	0	90	0	90	0	90	1	89
3		0	10	3	7	3	7	3	6	4	6
		0	90	0	90	1	90	1	90	2	88
5		0	10	6	4	6	4	7	3	8	2
		0	90	0	90	0	90	0	90	1	89
7		0	10	7	3	7	3	8	2	9	1
		0	90	0	90	0	90	0	90	1	89

NOTE: The table contains the average (over the Monte-Carlo replications) of the percentage of outliers that are (not) detected. A 'cell' in the table consists of four numbers, indicating

average percentage of observations **correctly** classified as outlier	average percentage of observations **incorrectly** classified as 'clean'
average percentage of observations **incorrectly** classified as outlier	average percentage of observations **correctly** classified as 'clean'

The simulation setup is described in the note to Table II. The percentage of additive outliers is 5.

TABLE V

COINTEGRATION TESTS FOR THE FINNISH/US EXCHANGE RATE DATA

$\nu=\infty$				$\nu=5$			
p	$k-r$			p	$k-r$		
	3	2	1		3	2	1
1	92.7”’	5.8	0.2	1	64.6***	10.5	1.6
$2^{a,s}$	69.1”’	6.1	0.0	$2^{a,s}$	19.2	2.2	0.3
3	51.1***	11.3	0.4	3	23.5	6.0	0.2
4	33.4”	12.0	0.0	4	22.2	8.0	0.0
5	29.9’	7.6	0.0	5	18.9	5.7	0.1
6	34.8”	9.1	0.5	6	28.2’	7.2	0.4
7	34.8**	9.0	0.2	7	31.7	4.5	0.2

The table contains the PLR (trace) tests for the hypothesis Ho : $r \leq 0$, H_1 : $r < 1$, and $H_2 : r \leq 2$, versus the alternative $H_3 : r = 3$. Here, r denotes the number of cointegrating relations. p denotes the order of the VAR model that is used for computing the test. ν denotes the degrees of freedom parameter that is used in the pseudo likelihood. *, **, and . ** denote significance at the 10, 5, and 1 per cent level, respectively. The critical values are taken from Table I for the case with drift, $\alpha'_{\perp}\mu \neq 0$. The superscripts ^{a} and ^{s} refer to the order of the model of the model that corresponds to the maximum value of the Akaike and the Schwarz information criterion, respectively.